Technical Report

–

Evaluation of children's sensitivity to central information of a narrative.

Jolien M. Mouw, Lesya Y. Ganushchak, & Paul van den Broek

Department of Education and Child Studies, Faculty of Social and Behavioral Sciences, Leiden University, Leiden.

The research reported in this technical report was made possible through the financial support of MUSE, EU FP7-296703.

1. Context

This evaluation is part of the Machine Understanding for interactive StorytElling (MUSE) project, a European collaborative project that aims to introduce a new way of exploring and understanding information by "bringing text to life" through 3D interactive storytelling. Taking as input natural language texts like children's stories, MUSE will process the natural language, translate it into formal knowledge that represents the actions, actors, plots, and surrounding world, and then render these as virtual 3D worlds in which the user can explore the text through interaction, re-enactment and guided game play. To enable such a system, MUSE aims to make targeted advances in natural language processing that enable the translation of natural language text to the necessary knowledge representations, as well as targeted advances in the action representation and story planning necessary for interactive storytelling. In interactive storytelling, MUSE develops action and object representations that bridge the gap between natural language and virtual worlds, and creates advanced techniques for planning virtual world stories given inconsistent and incomplete information.

One of the goals of the MUSE project is to increase the extent to which user-constructed representations match the intended knowledge representation. Therefore, as a first step, the nature of information that is included in these user-constructed representations was examined.

2. Evaluation of Central Information

In the fall of 2013, a pilot study was conducted to evaluate children's sensitivity to central information of our target narrative *The day Tuk became a hunter* (Melzack, 1968; from now we will refer to the narrative as Tuk the Hunter). For each user, a measure of sensitivity to structural centrality was taken (Kintsch, 1998; Van den Broek, 2010), resulting in a user-selected indication of importance of each sentence.

2.1 Sampling Procedure

Participants were fifth grade students from Dutch elementary schools. Primary schools were contacted by telephone and when a school expressed their willingness to participate,

they were selected for the study (convenience sampling). Parents and children were given written information about the study at forehand. In order for a child to participate in this study, parental agreement (active consent) was needed. During our study, children were allowed to stop at any time they wanted. All 142 participating children (60 girls and 82 boys; mean age of 10 years and 8 months) from 6 different schools finished the tasks. In return for their partaking, all participants received a small gift.

2.2 Materials

Reading comprehension and the user-constructed representations were measured by means of a reading task. Children were asked to first read (the Dutch version of) Tuk the Hunter. Next, they were instructed to "Mark the 10 sentences that you think are most important. For example, if you were to make a movie, what would be included?". This way, it was possible to examine readers' sensitivity to important and/or central text elements (i.e. mental representation) in the story Tuk the Hunter. The narrative had 66 sentences in total. All children finished this assignment within 30 minutes.

2.3 Data Preparation

Data for stories in which children marked 10 sentences were included in the analysis. Additionally, data for stories in which 9 or 11 sentences were marked was included as well. The decision was made because of aspects of the Dutch translation of the story (i.e., splitting/ merging sentences) and/or punctuation marks. For example, in the sentence "*No, Tuk, keep away!" shouted his father. "Don't come near!*", three punctuation marks were included. Some children only marked the first part, up to the exclamation mark. Some children reasoned that "shouted his father" was also part of the same sentence, whereas others marked the whole sentence. Thus, some children counted this sentence as constituting three markings, whereas others as one.

Data was excluded if more than 11 or fewer than 9 sentences were selected, as these children might have misunderstood the task. In addition, data was excluded if all marked sentences were on the first page, as this could be an indication that a child did not take into

account that key phrases could also be found in the second part of the story and/or did not read the story first before starting to mark sentences. Data from children with dyslexia were also excluded. Data of 130 cases remained for further analysis.

2.4 Results

In these analyses, we evaluated the data from three perspectives:

1. What sentences did the children actually select (i.e. the number of children that have selected a particular sentence)?
2. Did children select sentences with many causal connections? That is based on a causal analysis to determine the structural importance of the selected sentences.
3. Do the selected sentences include the story's main events?

For each of these perspectives to evaluate children's sensitivity to central information, we have plotted the frequency with which each sentence is selected against its number of causal connections (see section 2.4.2).

These results (Figure 1) also give valuable insights in children's' reasoning to be taken into account when designing evaluation instruments. For example, one child has selected the title as the most important sentence, whereas we did not include the sentence in our causal analysis.

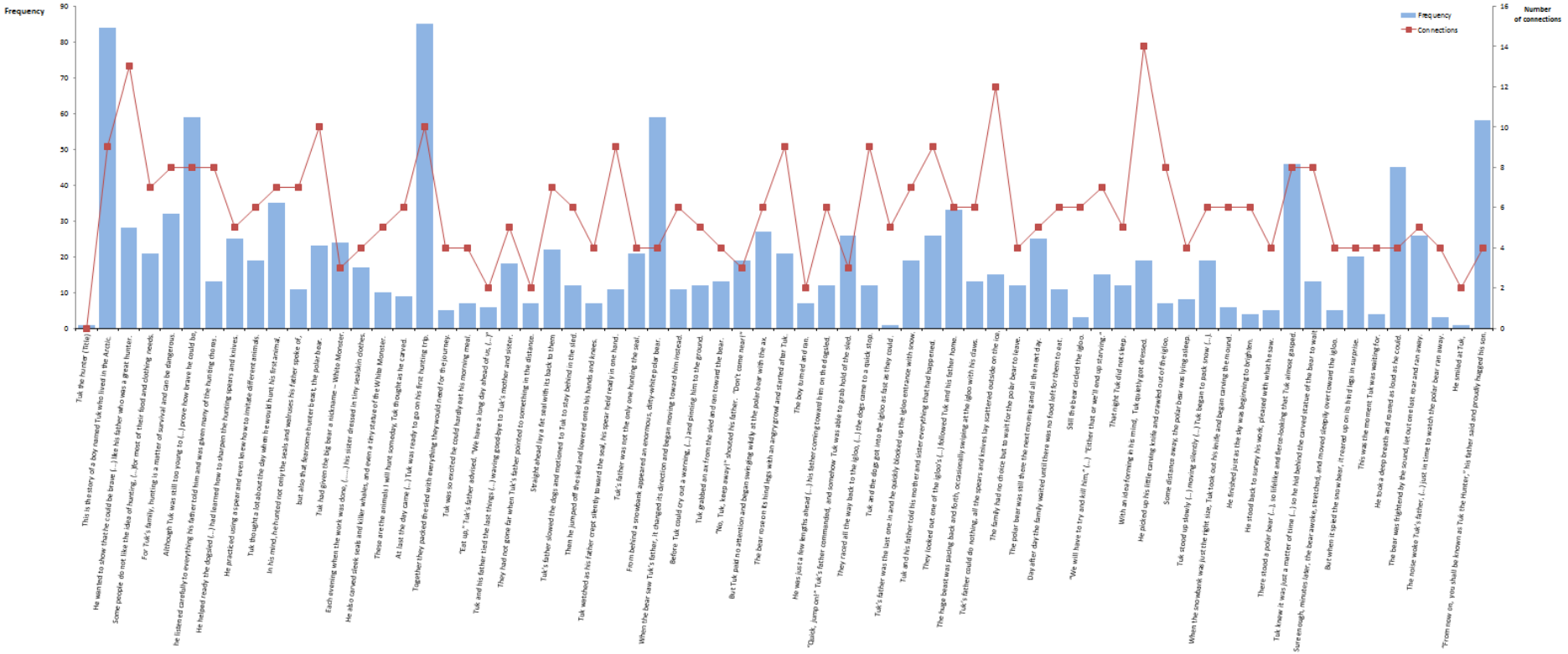


Figure 1. Overview of the whole story, with the data points on the x axis representing each of the 66 sentences of Tuk the Hunter. The bars represent the number of children that have selected a particular sentence, and the red line indicates the number of causal connections a sentence has.

2.4.1 Frequency of Selection

Figure 2 presents an overview of the 10 sentences that were most often selected (i.e. selected by most children). On the x axis the 10 most often selected sentences are presented. The left y axis represents the frequency in which this sentence was selected, whereas the right y axis represents the number of causal connections the selected sentence has. The sentence most often selected by the children is Sentence 16 (*At last the day came when his father announced that Tuk was ready to go on his first hunting trip*.). More than 65.3% of the children thought that this sentence was most important. However, this sentence only has 10 causal connections. The sentence that was rated as the second most important sentence, is Sentence 1 (*This is the story of a boy named Tuk who lived in the Arctic*.). It has been selected by 64.6% of the children, whereas it has only 9 connections. This implies that even though not all children agreed which were the most important sentences, most children did agree that these two sentences should be considered as important for story understanding.

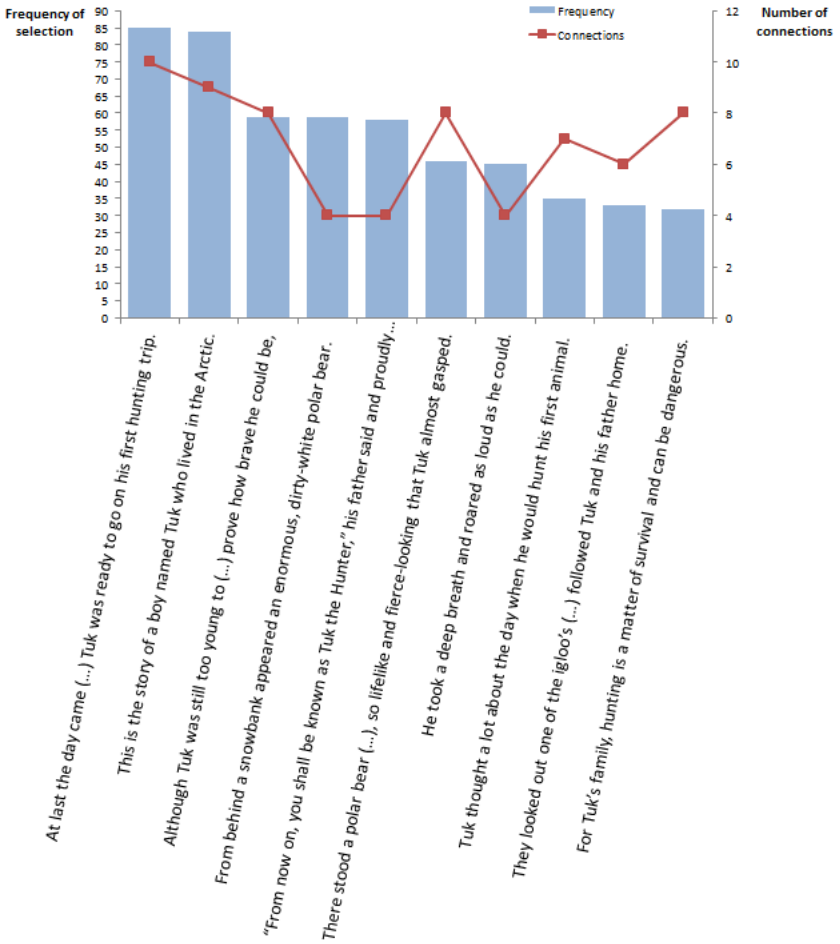
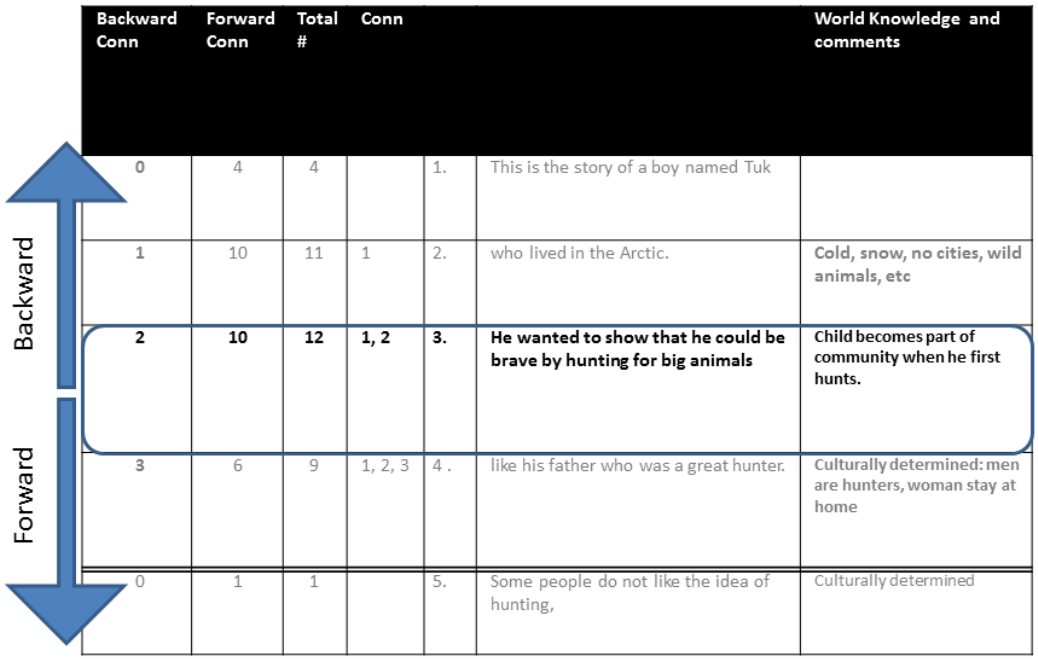


Figure 2. Overview of the ten sentences that are selected most often.

2.4.2 Causal Analysis

Next, it is examined if children indeed selected those sentences that are most central to the story, as we determined in our causal analysis. Following Van den Broek (1994), we identified two types of connections in our causal analysis, namely backward and forward connections (see Figure 3).



Backward Conn	Forward Conn	Total #	Conn			World Knowledge and comments
0	4	4		1.	This is the story of a boy named Tuk	
1	10	11	1	2.	who lived in the Arctic.	Cold, snow, no cities, wild animals, etc
2	**10**	**12**	**1, 2**	**3.**	**He wanted to show that he could be brave by hunting for big animals**	**Child becomes part of community when he first hunts.**
3	6	9	1, 2, 3	4 .	like his father who was a great hunter.	Culturally determined: men are hunters, woman stay at home
0	1	1		5.	Some people do not like the idea of hunting,	Culturally determined

Figure 3. Examples of forward and backward causal connections in Tuk the Hunter.

Backward connections refer to the information that was given prior to the target sentence (e.g. Sentence 3 in Figure 3). Forward connections are going forward towards the end of the story; it was also calculated how many time this sentence was required for the understanding of other sentences. To identify both types of connections, the following criteria were used:

1) What does the reader need to know from the previous sentences to understand the current one?

2) Always look for the closest possible connection to a sentence first. As a result, for example, not all sentences that contained the concepts Arctic and Tuk were connected to the first two sentences in which these concepts were introduced for the first time.

The causal analysis resulted in an inventory of all possible forward and backward connections that each sentence has. The total number of connections is the combined number of forward and backward connections.

Based on these calculations, we selected 10 out of 66 sentences that have the highest number of causal connections. Figure 4 presents an overview of the 10 sentences with the largest number of causal connections (red line), plotted against the frequency with which children selected these sentences (blue bars). According to our selection, Sentence 50 (*With an idea forming in his mind, Tuk quietly got dressed*.) is the most central sentence, as it has 14 forward and backward connections. However, only 14.6% of the children selected this as one of the 10 most important sentences. The second most central sentence, "*He wanted to show that he could be brave by hunting for big animals like his father who was a great hunter*", has 13 connections and was selected only by 21.5% of the children. Overall, only two of the most selected sentences were also in the top 10 most central sentences (ranked 5th and 6th, with 10 and 9 connections, respectively).

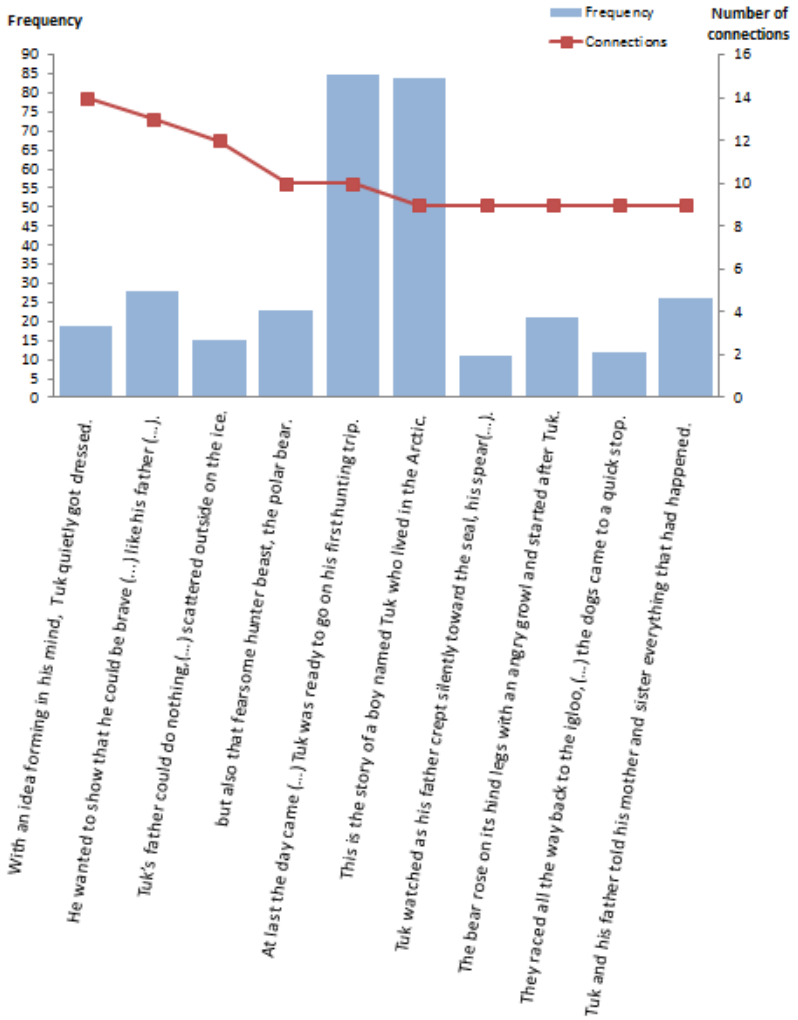


Figure 4. Overview of the ten sentences with most causal connections.

2.4.3 Main Events

Before conducting the study, we selected 12 sentences from the story that can be considered as the main events (e.g., expert agreement on 12 main events that summarize the story). The criteria used to select the sentences was that these sentences were essential for the unwinding of the story by considering what information, which actions, and/or what characters are needed in order for the story to end the way it does. Figure 5 depicts the 12 sentences that were considered as the main events. If a reader would read these 12 sentences consecutively, he or she would receive an accurate summary of the story and should be able to understand the main course of events.

We examined whether the 10 sentences the children selected most often corresponded to these 12 main events. The more main events a child selected, the better we judged his/her story comprehension. On average, children selected 4.18 (*SD* = 1.69, range; minimum = 0, maximum = 8) of the twelve main events.

In conclusion, these main events include both the most central sentence, as decided on based on the causal analysis, and the most selected sentences. For example, in Figure 5, one sentence is marked with an orange dot (*With an idea forming in his mind, Tuk quietly got dressed*.). This is the sentence with the highest number of causal connections. The bar, highlighted with the orange border, represents the most frequently selected sentence ("*At last the day came when his father announced that Tuk was ready to go on his first hunting trip*").

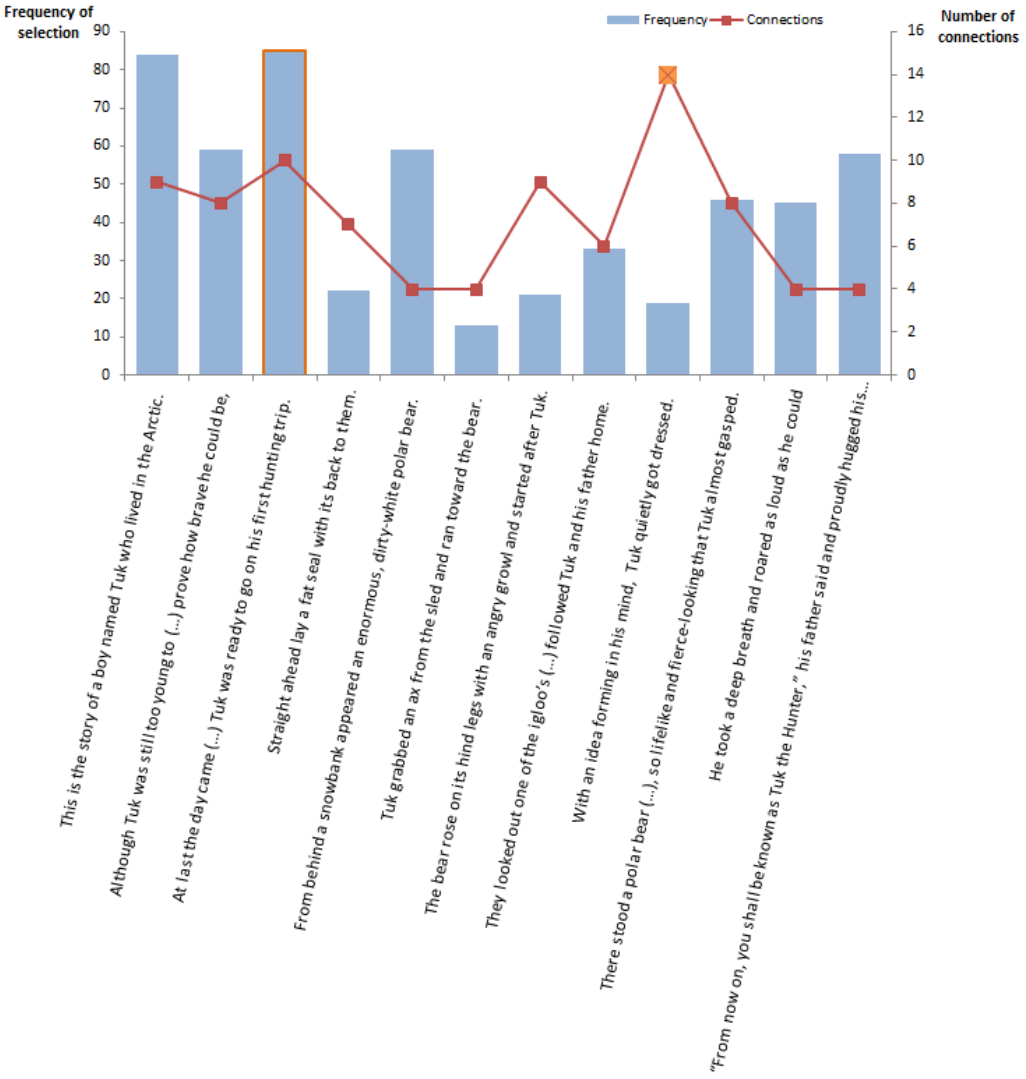


Figure 5. Overview of 12 sentences selected as main event.

3. Implications for Hypotheses Testing

The aim of the MUSE project is to increase the extent to which user-constructed representations matches the intended knowledge representation. Therefore, it first needed to be examined what information is included in these user-constructed representations. This pilot study shows that the sentences that are most central (both the most causally connected and the main events) in the story were not most often selected by the children. It seems that the user-constructed representation not necessarily corresponds to the intended knowledge representation. Therefore, these results help to generate hypotheses about the role of causal structure and knowledge representation in children's processing of the MUSE version of Tuk the Hunter (i.e., the 3D visualization of the story). In addition, these findings will be used to

determine situations where interactivity might help improve children's sensitivity to the structure of the narrative.

4. References

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York, NY: Cambridge University Press.

Melzack, R. (1968). *The day Tuk became a hunter and other Eskimo stories*. London: Macmillan.

Van den Broek, P. W. (1994) Comprehension and memory of narrative texts: Inferences and coherence. In: Gernsbacher, M.A. (Ed.), *Handbook of psycholinguistics*, 539-588. New York, NY: Academic Press.

Van den Broek, P. W. (2010). Using texts in science education: Cognitive processes and knowledge representation. *Science, 328*, 453-456.